SFB 649 Discussion Paper 2010-029

Adaptive Interest Rate Modelling

Mengmeng Guo*
Wolfgang Karl Härdle*

* Humboldt Universität zu Berlin, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

SFB 649 ECONOMIC RISK BERLIN

Adaptive Interest Rate Modelling*

Mengmeng Guo[†], Wolfgang Karl Härdle [‡]

May 21, 2010

Abstract

A good description of the dynamics of interest rates is crucial to price derivatives and to hedge corresponding risk. Interest rate modelling in an unstable macroeconomic context motivates one factor models with time varying parameters. In this paper, the local parameter approach is introduced to adaptively estimate interest rate models. This method can be generally used in time varying coefficient parametric models. It is used not only to detect the jumps and structural breaks, but also to choose the largest time homogeneous interval for each time point, such that in this interval, the coefficients are statistically constant. We use this adaptive approach and apply it in simulations and real data. Using the three month treasure bill rate as a proxy of the short rate, we find that our method can detect both structural changes and stable intervals for homogeneous modelling of the interest rate process. In more unstable macroeconomy periods, the time homogeneous interval can not last long. Furthermore, our approach performs well in long horizon forecasting.

Keywords: CIR model, Interest rate, Local parametric approach, Time homogeneous interval, Adaptive statistical techniques.

JEL classification: E44, G12, G32, N22

---

*The financial support from the Deutsche Forschungsgemeinschaft via SFB 649 "Ökonomisches Risiko", Humboldt-Universität zu Berlin is gratefully acknowledged.

†Research Associate at the Institute for Statistics and Econometrics of Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin, Germany and Wang Yanan Institute for Studies in Economics (WISE), Xiamen University, Xiamen, China. Email: guomengm@cms.hu-berlin.de;

‡Professor at Humboldt-Universität zu Berlin and Director of C.A.S.E - Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin, Germany and National Central University, Department of Finance, No. 300, Jhongda Rd., Jhongli City, Taoyuan County 32001, Taiwan (R.O.C.). Email: haerdle@wiwi.hu-berlin.de

1 Introduction

Interest rate risk, caused by the variability of interest rates, is the risk borne by an interest-bearing asset, such as a loan or a bond. It is commonly measured by the bond's duration. Interest rate risk causes credit risk, which in turn may induce further risks (subprime crisis). For hedging purposes, however, it is important to price the interest rate derivatives which of course depend on the dynamic process of the interest rate. If the macroeconomy is unstable, the variation of interest rate will be larger, and vice versa. For instance, in 2002, bubbles existed in the stock market. In 2003, the war in Iraq influenced the macroeconomy. Since 2007, the macroeconomy has been depressed by the subprime crisis. We find the short rate in these periods more fluctuating. On the other hand, changes in business cycle conditions may influence the dynamics, and they may differ from one period to another. The stochastics of short rates are impacted by these facts, which return shock reverting behaviour of stochastic interest rate volatility less persistent to economic shocks. These shocks or news are dominated by central bank announcements of base rate changes. The short rates respond quickly to these unanticipated announcements. This conclusion can be supported by Jones, Lmont, and Lumsdaine (1998) who documented that volatility shocks to U.S. treasure bonds arising from scheduled macroeconomic announcements are not persistent at all. A large number of empirical studies have demonstrated the unstable property of the interest rate process. The instability can be induced by structural breaks, or regime switchings, even the process can be a smooth function of time.

Due to this instability in statistical modelling, a wide variety of interest rate models have been introduced. Three main strands of literature exist to describe the dynamics. On one hand, the instability is induced by the structural breaks, which are captured by jump diffusion models. In this kind of model, it is assumed that several jumps exist in the diffusion function to capture the structural breaks. Das (2002) incorporated jumps into the Vasicek model and found strong evidence of jumps in the daily federal funds rate. Johannes (2003) used a nonparametric diffusion model to study the secondary three month treasury bills. He concluded that jumps are generally generated by the arrival of news about the macroeconomy. A common conclusion is that the nonlinearity exists in the dynamics of short rates. Another strand of literature uses regime switching models to capture the business cycle character of interest rates, see Ang and Bekaert (2002), Bansal and Zhou (2002). They found that the interest rate has significant changes and its variation performs differently in different regimes. In the third kind of model, the process parameters (drift or volatility) are assumed to be a function of time, Hull-White (1990), Black-Karasinski (1991), Aït-Sahalia (1996) and Stanton (1997), Fan et al (2003) and Arapis and Gao (2006). The conclusion of them is that the coefficients in the CIR model are time varying, moreover they proved that the drift function is nonlinear. Generally speaking, a one factor short rate model with constant parameters may not be valid for a long period of time in reality.

In this paper, we introduce a time varying CIR model and estimate it from a novel point of view - the local parametric approach (LPA). Our aim is to find the longest stable "time homogeneous" interval for each time point t, where the parameters in the CIR model can be safely assumed to be constant. More attractively, by this method, we can detect jumps and structural break points. Furthermore, this approach includes regime switching models and it also describes the time variation of coefficients. Based on the parameters inside the selected interval, one may distinguish blooming and declining regimes of the economy.

The proposed approach may be applied to different problems. Giacomini, Härdle and Spokoiny (2009) considered time varying copulae estimation, Čížek, Härdle and Spokoiny (2009) applied it to compare the performance of global and time varying ARCH and GARCH specifications, Härdle, Okhrin and Okhrin (2010) applied this method to hierarchical archimedean copulae, and found that the LPA can be used to detect both adaptive copulae parameters and local dependency structures.

To assess the performance of the LPA, we do both simulations and empirical studies. In the simulation exercise, we show that the proposed LPA detects the structural breaks very well, and the true parameters are located in the pointwise confidence intervals of the estimators. In the empirical study, we use three month treasure bill rate as a proxy of the short rate and investigate the performance of the LPA by both in sample fitting and out of sample forecasting via comparing with moving window estimators.

The rest of this paper is organized as follows. In Section 2, we give a short recall about one factor interest rate models, later we explain the LPA in detail in Section 3. In Section 4, we present our simulation results. Empirical studies are presented in Section 5. We conclude in Section 6.

2 Interest Rate Models

We give a short review about one factor interest rate models. Note that we need the essential properties:

- Mean reversion: the interests rates always tend to return to an average level.
- The interest rate $r(t)$ is non negative.

Vasicek Model

$$dr(t) = a\{b - r(t)\}dt + \sigma dW_t$$

where a, b and σ are constants, W_t is a standard Brownian process. It is consistent with the mean reversion feature with a reversion speed a to the long run mean level b. However, in this model $r(t)$ can be negative.

Cox, Ingersoll and Ross (CIR) Model

$$dr(t) = a\{b - r(t)\}dt + \sigma\sqrt{r(t)}dW_t \tag{1}$$

The drift function $\mu\{r(t)\} = a\{b - r(t)\}$ is linear and possesses a mean reverting property, i.e. $r(t)$ moves in the direction of its long run mean b at speed a. The diffusion function $\sigma^2\{r(t)\} = r(t)\sigma^2$ is proportional to the interest rate $r(t)$ and ensures that the process stays on a positive domain. Here $r(t)$ has a positive impact on the standard deviation through (1).

Hull-White Model

$$dr(t) = \{\delta(t) - ar(t)\}dt + \sigma dW_t$$

This is an extended Vasicek model, where a and σ are constant, $\delta(t)$ is a deterministic function of time. Moreover, this model uses the time dependent reversion level $\delta(t)/a$ instead of the constant b in Vasicek model.

Black-Karasinski Model

$$d\log r(t) = \delta(t)\{\log\mu(t) - \log r(t)\}dt + \sigma(t)dW_t$$

with $\delta(t)$, $\mu(t)$ and $\sigma(t)$ as a deterministic function of time, where $\mu(t)$ as the target interest rate. A drawback of this model is that no closed form formula for valuing bonds in terms of $r(t)$ can be derived by this model.

3 Methodology

In the Vasicek model, the interest rate $r(t)$ can be negative. As an improvement of the Vasicek model, the CIR model guarantees the interest rate is not negative. In the Hull-White model, the volatility is a constant. The Black-Karasinski model assumes $\delta(t)$ and $\mu(t)$ are deterministic function of time. However, all these models are too restrictive due to the unstable macroeconomy. In this section, a new method - the LPA for time varying CIR model is introduced. The method allows the parameters vary over time as the time homogeneous interval changes with t. Discontinuities and jumps may be detected and used to identify structural changes.

The time varying CIR model is expressed as:

$$dr(t) = a_t\{b_t - r(t)\}dt + \sigma_t\sqrt{r(t)}dW_t \tag{2}$$

where, W_t is the standard Wiener Process. Denote the time varying parameters as $\theta_t = (a_t, b_t, \sigma_t)^\top$. This CIR model (2) includes all of the aforementioned parametric models, such as jump diffusion models, regime switching models, and also the non-parametric specified time varying interest rate models.

The discrete version of (2) is:

$$Y_i = r_{t_{i+1}} - r_{t_i} = a_t\{b_t - r_{t_i}\}\Delta t + \sigma_t\sqrt{r_{t_i}}Z_i \tag{3}$$

Where $\{Z_i\}_{i=1}^T$ are normally distributed with zero mean and variance $\Delta t = t_{i+1} - t_i$, (more generally, Z_i can be a white noise process). The time unit may be one year, $\Delta t = \frac{1}{250}$ for daily data, or for weekly data, $\Delta t = \frac{1}{52}$.

3.1 Likelihood Function of CIR Process

If a, b, σ are all positive, and $2ab \geq \sigma^2$ holds, then the CIR model is well defined and has a steady state distribution. Given r_t at time t, the density of $r_{t+\Delta t}$ at time point $t + \Delta t$ is:

$$p(r_{t+\Delta t}|r_t; \theta, \Delta t) = ce^{-u-v}(\frac{v}{u})^{\frac{q}{2}}I_q(a\sqrt{uv}) \tag{4}$$

where

$$
\begin{aligned}
c &= \frac{2a}{\sigma^2(1-e^{-a\Delta t})} \\
u &= cr_t e^{-a\Delta t} \\
v &= cr_{t+\Delta t} \\
q &= \frac{2ab}{\sigma^2}
\end{aligned}
$$

and $I_q(2\sqrt{uv})$ is the modified Bessel function of the first kind with order q. The log likelihood function is given by:

$$
L(\theta) = \sum_{i=1}^{T-1} \log p(r_{t_{i+1}}|r_{t_i};\theta,\Delta t) \tag{5}
$$

Fix now t, the MLE estimator $\tilde{\theta}_{I_k}$ in any interval $I_k = [t-m_k, t]$ is:

$$
\tilde{\theta}_{I_k} = \arg\max L_{I_k}(\theta) = \arg\max \sum_{i\in I_k} \log p(r_{t_{i+1}}|r_{t_i};\theta,\Delta t)
$$

The accuracy of the estimation for a locally constant model with parameter θ_0 is measured via the log likelihood ratio $L_{I_k}(\tilde{\theta}_{I_k},\theta_0) = L_{I_k}(\tilde{\theta}_{I_k}) - L_{I_k}(\theta_0)$. In Čížek, Härdle and Spokoiny (2009), it is proved that if Y_i follows a nonlinear process (2), then given I_k for any $r > 0$, there exists a constant $\Re_r(\theta_0)$, such that:

$$
\mathsf{E}_{\theta_0} \left|L_{I_k}(\tilde{\theta}_{I_k},\theta_0)\right|^r \leq \Re_r(\theta_0) \tag{6}
$$

Thus, $\Re_r(\theta_0)$ can be treated as the parametric risk bound. It enables testing the parametric hypothesis on the basis of the fitted log likelihood $L_{I_k}(\tilde{\theta}_{I_k},\theta_0)$.

3.2 Test of Homogeneous Intervals

Mercurio and Spokoiny (2004), Čížek, Härdle and Spokoiny (2009) and Spokoiny (2009) are informative references for the LPA. The general idea can be described as follows: suppose we have K (historical) candidate intervals with a starting interval I_0, i.e. $I_0 \subset I_1 \subset \cdots \subset I_K$, $I_k = [t-m_k, t]$ with $0 < m_k < t$. We increase the length from m_k to m_{k+1}, and test over the larger interval I_{k+1} whether $\tilde{\theta}_{k+1}$ is still consistent with $\tilde{\theta}_k$. To test an interval $I_k = [t-m_k, t]$, we set the null (2) with a fixed parameter θ. The alternative is to find an unknown change point τ in the interval I_k, i.e. Y_l follows one process when $t' \in J = [\tau+1, t]$ with parameter θ_J, and it follows another process when $t' \in J^c = [t-m_{k+1}, \tau]$ with parameter θ_{J^c}, where $\theta_J \neq \theta_{J^c}$.

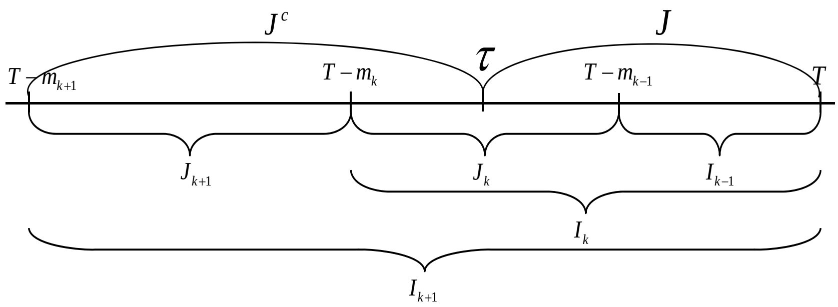

Figure 1: Construction of the Test Statistics for LPA: the involved interval I_k and J_k.

With this alternative, the likelihood can be expressed as $L_J(\tilde{\theta}_J) + L_{J^c}(\tilde{\theta}_{J^c})$, giving the test statistics:

$$T_{I_{k+1},\tau} = L_J(\tilde{\theta}_J) + L_{J^c}(\tilde{\theta}_{J^c}) - L_{I_{k+1}}(\tilde{\theta}_{I_{k+1}}) \tag{7}$$

where $\tau \in J_k = I_k \backslash I_{k-1}$, see Figure 1. Since the change point $\tau \in I_k$ is unknown, we consider the maximum of the test statistics over J_k:

$$T_k = \max_{\tau \in J_k} T_{I_{k+1},\tau} \tag{8}$$

The selected longest time homogeneous interval satisfies

$$T_k \leq \mathfrak{z}_k, \qquad \text{for} \quad k \leq \hat{k} \tag{9}$$

and $T_{\hat{k}+1} > \mathfrak{z}_{\hat{k}+1}$. Then $I_{\hat{k}}$ is the longest time homogeneous interval for time point t, and the local adaptive estimator $\hat{\theta}_t = \hat{\theta}_{I_{\hat{k}}}$. The event $\{I_k \text{ is rejected}\}$ means that $T_\ell > \mathfrak{z}_\ell$ for some $\ell < k$, and hence a change point has been detected in the first k steps within I_k. $\mathfrak{z}_k$ is the critical value (depending on the interval sequence I_k) to be introduced later.

3.3 The Local Parametric Approach (LPA)

For any given t with the intervals $I_0 \subset I_1 \subset \cdots \subset I_K$, following the idea mentioned above, the algorithm is described into four steps.

1. We estimate $\tilde{\theta}_{I_0}$ using the observations from the smallest interval $I_0 = [t - m_0, t]$, $\tilde{\theta}_{I_0}$ is always accepted.

2. We increase the interval to $I_k, (k \geq 1)$, get the estimator $\tilde{\theta}_{I_k}$ by MLE, and test homogeneity via (7), i.e. we test whether there is a change point in I_k. If (9) is fulfilled, we go on to step 3, otherwise we go to step 4.

3. Let $\hat{\theta}_{I_k} = \tilde{\theta}_{I_k}$, then further set $k = k+1$, and go to step 2.

4. Accept as the longest time homogeneous interval $I_{\hat{k}} = I_{k-1}$, and define the local adaptive estimator as $\hat{\theta}_{I_{\hat{k}}} = \tilde{\theta}_{I_{k-1}}$. Additionally set $\hat{\theta}_{I_{\hat{k}}} = \hat{\theta}_{I_k} = \cdots = \hat{\theta}_{I_K}$ for all $k > \hat{k}$.

For a change point τ in I_k, we obtain $\hat{k} = k-1$, and $I_{\hat{k}} = I_{k-1}$ is the selected longest time homogenous interval. We compare the test statistics with the critical values, if it is smaller than the critical value $\mathfrak{z}_k$, we accept I_k as the time homogeneous interval, then we increase the interval to I_{k+1}, and do the test again. We sequentially repeat this procedure until we stop at some $k < K$ or we exhaust all the chosen intervals. For each time point t, we use the same algorithm, and we do not to calculate the critical values a second time, since they depend on only the parametric specification and the length of interval m_k, not on t.

To investigate the performance of the adaptive estimator, we introduce the small modelling bias (SMB). The SMB for interval I_k is:

$$\Delta_{I_k}(\theta) = \sum_{t \in I_k} \mathcal{K}\{r(t), r(t;\theta)\} \tag{10}$$

with $\mathcal{K}$ the Kullback-Leibler (KL) divergence,

$$\mathcal{K}\{r(t), r(t;\theta)\} = \mathsf{E} \log \frac{p\{r(t)\}}{p\{r(t;\theta)\}} \tag{11}$$

where $p(.)$ and $p(.;\theta)$ are the probability density functions of $r(t)$ and $r(t;\theta)$ respectively. The SMB measures in terms of KL divergence the closeness of a constant parametric model with $p(.;\theta)$ to a time varying model with $p(.)$. Suppose now that for a fixed $\Delta > 0$:

$$\mathsf{E}\, \Delta_{I_k(\theta)} \leq \Delta \tag{12}$$

(12) simply means that for some $\theta \in \Theta$, $\Delta_{I_k}(\theta)$ is bounded by a small constant with a high probability, which implies the time varying model can be well approximated over I_k by the parametric model with fixed parameter θ.

Under this SMB condition for some interval I_k and $\theta \in \Theta$, one has the property:

$$\mathsf{E} \log\{1 + \frac{|L_{I_k}(\tilde{\theta}_{I_k}, \theta)|^r}{\mathfrak{R}_r(\theta)}\} \leq 1 + \Delta \tag{13}$$

If Δ is not large, (13) extends the parametric risk bounds to the nonparametric situation under the SMB condition, see Čížek, Härdle and Spokoiny (2009). An "oracle" choice I_{k^*} from the set $I_0, \cdots, I_K$ exists, which is defined as the largest interval satisfying (12). We denote the corresponding "oracle" parameter as $\theta_{I_{k^*}}$.

However, two types of errors occur in this algorithm: the first type is to reject the time homogeneous interval earlier than the "oracle" step, which means $\hat{k} \leq k^*$. The other type is to reject the homogeneous interval later than the "oracle" step, i.e. $\hat{k} > k^*$. The first type of error can be treated as a "false alarm", i.e. the algorithm stops earlier than the "oracle" interval I_{k^*}, which leads to selecting an estimate with a larger variation than $\theta_{I_{k^*}}$. The second type of the error arises if $\hat{k} > k^*$. Outside the oracle interval we are exploiting data which does not support the SMB condition. Both errors will be specified in a propagation and stability condition in the next section.

3.4 Choice of Critical Values

The accuracy of the estimator can be measured by the log likelihood ratio $L_{I_k}(\tilde{\theta}_{I_k}, \theta_0)$ between the MLE estimator and the true parameters in parametric specification, which is stochastically bounded by the exponential moments (13). In general, $\tilde{\theta}_{I_k}$ differs from $\hat{\theta}_{I_k}$ only if a change point is detected at the first k steps. A small value of the likelihood ratio means that $\hat{\theta}_{I_k}$ belongs to the confidence set based on the estimate of $\tilde{\theta}_{I_k}$, i.e. statistically we accept $\hat{\theta}_{I_k} = \tilde{\theta}_{I_k}$. If the procedure stops at some $k \leq K$ by a false alarm, i.e. a change point is detected in interval I_k with the adaptive estimator $\hat{\theta}_{I_k}$, then the accuracy of the estimator can be expressed

$$\mathsf{E}_{\theta_0} \left| L_{I_k}(\tilde{\theta}_{I_k}, \hat{\theta}_{I_k}) \right|^r \leq \rho \mathfrak{R}_r(\theta_0) \tag{14}$$

which is to be referred as the "propagation" condition. We choose the critical value $\mathfrak{z}_l$ based on (14). The situation at the first k steps can be distinguished into two cases. One is that a change point is detected at some step $l \leq k$, otherwise there is no change point in the first k intervals. We denote by $\mathfrak{B}_l$ the event of rejection at step l, that is,

$$\mathfrak{B}_l = \{T_1 \leq \mathfrak{z}_1, \cdots, T_{l-1} \leq \mathfrak{z}_{l-1}, T_l > \mathfrak{z}_l\} \tag{15}$$

and $\hat{\theta}_{I_k} = \tilde{\theta}_{I_{l-1}}$ on $\mathfrak{B}_l$, $l = 1, 2, \cdots, k$. We choose $\mathfrak{z}_1$ by minimizing the following equation:

$$\max_{k=1,\cdots,K} \mathsf{E}_{\theta_0} \left| L(\tilde{\theta}_{I_k}, \tilde{\theta}_{I_0}) \right|^r \mathbf{1}(\mathfrak{B}_1) \leq \rho \mathfrak{R}_r(\theta_0)/K$$

For $\mathfrak{z}_l, l \geq 2$, We use the same algorithm to calculate them. The event $\mathfrak{B}_l$ only depends on $\mathfrak{z}_1, \cdots, \mathfrak{z}_l$. Since $\mathfrak{z}_1, \cdots, \mathfrak{z}_{l-1}$ have been fixed by previous steps, the event $\mathfrak{B}_l$ is controlled by $\mathfrak{z}_l$. Hence, the minimal value of $\mathfrak{z}_l$ should ensure

$$\max_{k \geq l} \mathsf{E}_{\theta_0} \left| m_k \mathcal{K}(\tilde{\theta}_k, \tilde{\theta}_{l-1}) \right|^r \mathbf{1}(\mathfrak{B}_l) = \rho \mathfrak{R}_r(\theta_0)/K \tag{16}$$

or we can express the criteria via the log likelihood ratio:

$$\max_{k \geq l} \mathsf{E}_{\theta_0} \left| L(\tilde{\theta}_{I_k}, \tilde{\theta}_{I_{l-1}}) \right|^r \mathbf{1}(\mathfrak{B}_l) = \rho \mathfrak{R}_r(\theta_0)/K \tag{17}$$

where ρ and r are two global parameters, and m_k denotes the number of points in I_k. The role of ρ is similar to the level of the test in the hypothesis testing problem, while r describes the power of the loss function. We apply $r = 1/2$ in both the simulation and the real data analysis, since it makes the procedure more stable and robust against outliers. We also choose $\rho = 0.2$, however other values in the range $[0.1, 1]$ leads to similar results, see Spokoiny (2009).

The critical value $\mathfrak{z}_l$ which satisfies (17) can be found numerically by Monte Carlo simulations from the parametric model. It is a decreasing function with respect to the log length of interval. When the interval is small, it is easier to accept it as the time homogeneous interval, since there are not many jumps due to the short interval, while if we increase the length of interval, as the sample size increases, it contains more uncertain information, especially when big jumps or visible structural changes exist in the interval, therefore it tends to reject the homogeneous interval test statistics for larger interval, correspondingly the critical value should decrease.

The length of the intervals is assumed to geometrically increase with $m_k = [m_0 a^k]$. m_0 is the initial length of I_0, which is time homogeneous as default. a can be chosen from 1.1 to 1.3. However, the experiments reveal that the results are not sensitive to the choice of a. In the time varying CIR model, three parameters need to be estimated. To guarantee a reasonable quality of the estimation, large sample size is required. Therefore, we choose the length of the initial interval I_0 as $m_0 = 40$ and also choose $a = 1.25$. As already discussed, the interest rates are influenced by macroeconomic variables, and may also be subject to regime shifts. Therefore the longest interval we choose should cover one regime, and at least one change point will exist between the expansion and recession regimes. Referring to a business cycle of around 4 years, we choose the number of intervals $K = 15$, so that $m_K = 1136$ is the longest tested time homogeneous interval used in both simulation and empirical exercises in this paper.

3.5 "Oracle" Property of The Estimators

In this section, we discuss the "oracle" properties of the LPA estimators. Recall that for the "oracle" choice k^*, (12) holds, and it also holds for every $k \leq k^*$, while it does not hold for any $k \geq k^*$. The "oracle" choice I_{k^*} and $\theta_{I_{k^*}}$ are of course unknown. The LPA algorithm tries to mimic this oracle value. In Čížek, Härdle and Spokoiny (2009), it is proved that under the SMB condition, the "oracle" property of the LPA estimator $\hat{\theta}_{I_{\hat{k}}}$ satisfies the following property:

For $\theta \in \Theta$ and let $\max_{k \leq k^*} \mathsf{E}\,|L(\tilde{\theta}_{I_{k^*}}, \theta)|^r \mathbf{1}(\mathfrak{B}_1) \leq \mathfrak{R}_r(\theta)$, one has:

$$\mathsf{E}\log\{1 + \frac{|L_{I_{k^*}}(\tilde{\theta}_{I_{k^*}}, \theta)|^r}{\mathfrak{R}_r(\theta)}\} \leq 1 + \Delta \tag{18}$$

Further, we can obtain

$$\mathsf{E}\log\{1 + \frac{|L_{I_{k^*}}(\tilde{\theta}_{I_{k^*}}, \hat{\theta}_{I_{\hat{k}})|^r}}{\mathfrak{R}_r(\theta)}\} \leq \rho + \Delta \tag{19}$$

This theorem tells us that although the false alarm occurs before the "oracle" choice, i.e. $\hat{k} \leq k^*$, under the SMB condition, the adaptive estimator $\hat{\theta}_{I_{\hat{k}}}$ does not go far from the oracle value, which implies the LPA estimator does not induce large errors into the estimations.

The SMB condition doesn't hold if $\hat{k} > k^*$, which means the detected interval is bigger than the "oracle" interval. However, the LPA estimator $\hat{\theta}_{I_{\hat{k}}}$ satisfies Theorem 4.3 in Čížek, Härdle and Spokoiny (2009):

Let $\mathsf{E}\,\Delta_{I_{k^*}(\theta)} \leq \Delta$ for $k^* \leq K$, then $L_{I_{k^*}}(\tilde{\theta}_{I_{k^*}}, \hat{\theta})\mathbf{1}(\hat{k} \geq k^*) \leq \mathfrak{z}_{k^*}$,

$$\mathsf{E}\log\{1 + \frac{|L_{I_{k^*}}(\tilde{\theta}_{I_{k^*}}, \hat{\theta}_{I_{\hat{k}}})|^r}{\mathfrak{R}_r(\theta)}\} \leq \rho + \Delta + \log\{1 + \frac{\mathfrak{z}_{k^*}^r}{\mathfrak{R}_r(\theta)}\} \tag{20}$$

It means $\hat{\theta}_{I_{\hat{k}}}$ belongs with a high probability to the confidence interval of the oracle estimate $\tilde{\theta}_{I_{k^*}}$, i.e. it is still a reliable approximation for the oracle value $\theta_{I_{k^*}}$.

4 Simulation Study

We evaluate the performance of the LPA for the CIR model via simulations first. We simultaneously change all three parameters $(a_t, b_t, \sigma_t)^\top$ and assume there are

t	a	b	σ
$t \in [1, 500]$	0.2	0.04	0.03
$t \in [501, 1000]$	0.5	0.06	0.1
$t \in [1001, 1500]$	0.8	0.01	0.07

Table 1: The parameter settings for simulations of the CIR process

two change points for each parameter in the process. Further, the structural breaks occur at the same time in all three parameters. We simulate the CIR path 100 times with the sample size $T = 1500$. Table 1 summarizes the parameter settings for the simulations of the CIR model, the chosen values are in the range of parameters from the globally CIR estimators.

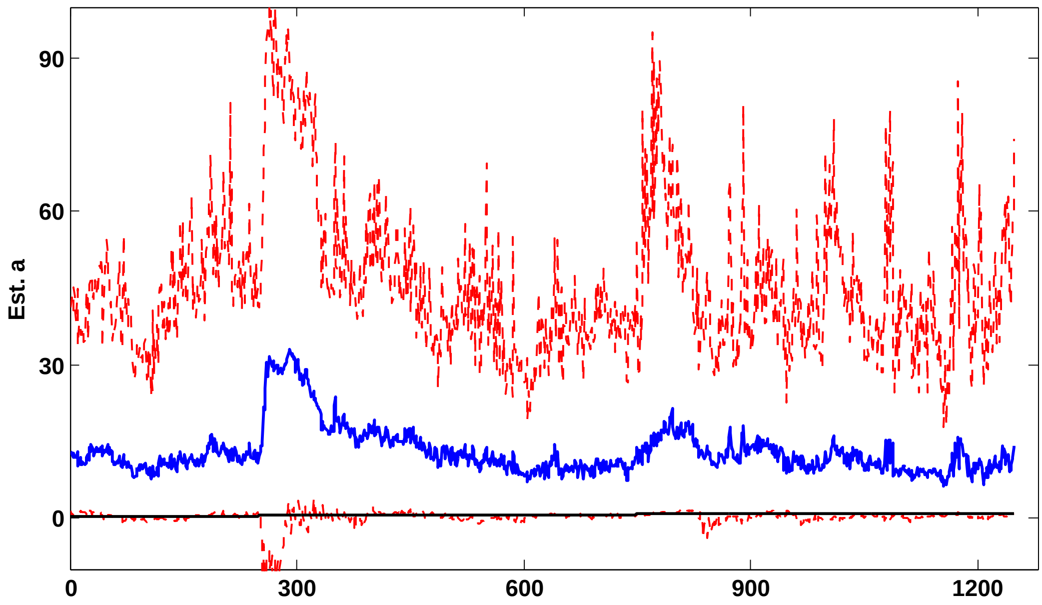

Figure 2: LPA estimator $\hat{a}$ with simulated CIR paths. The dotted red lines are the 5%–95% pointwise confidence intervals of $\hat{a}$, the blue line is the mean of $\hat{a}$, and the black line stands for the true process as set in Table 1.

The estimators $\hat{a}$, $\hat{b}$, and $\hat{\sigma}$ are described in Figures 2 to 5. The blue lines respectively depict the means of the corresponding estimators from the 100 simulations, and the two dotted red lines are the 5%–95% pointwise confidence intervals for the estimators. The black lines describe the respective real parameters. We use the first 250 data points as the training set referring to the moving window estimator, then we estimate the CIR model by the LPA from the time point 251 to 1500. One can

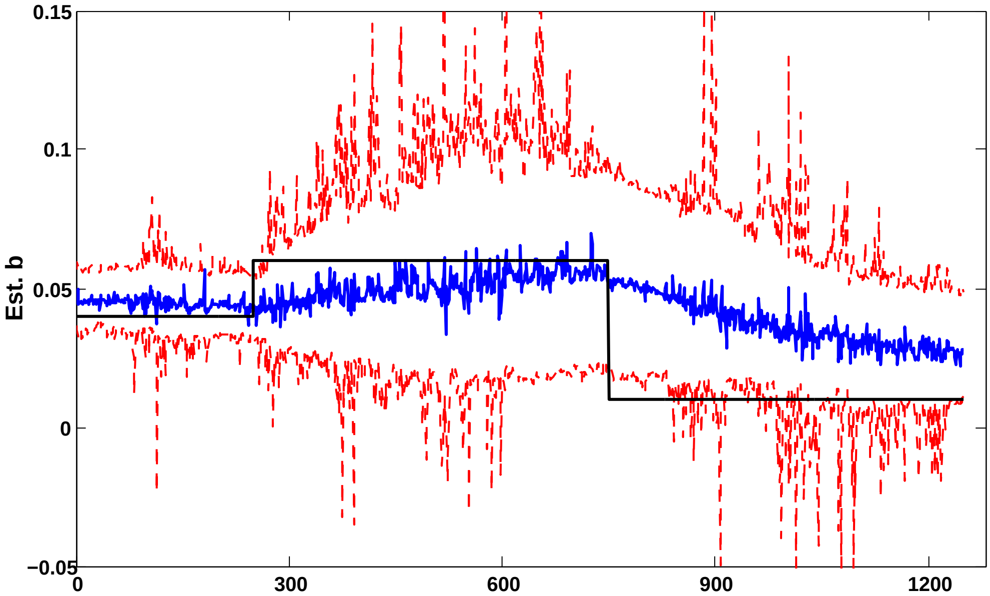

Figure 3: LPA estimator $\hat{b}$ with simulated CIR paths.The dotted red lines are the 5%–95% confidence interval of $\hat{b}$, the blue line is the mean of $\hat{b}$, and the black line stands for the true process as set in Table 1.

see that for the mean reversion speed a, the LPA under the null contains the true parameter.

Figure 3 presents the performance of the LPA estimator $\hat{b}$. Its performance is reasonable. One can obviously detect there are two jump points, which respectively locate around time point 300 and 800. Taking the delay time into consideration, the performance of $\hat{b}$ coincides with the true process.

It is worth noting that the performance of the LPA estimator $\hat{\sigma}$ is preferable to that of $\hat{a}$ and $\hat{b}$. The structural break points is obvious in Figure 4. Both the mean value and the confidence intervals have the same trend as the true parameter path, which indicates the LPA can capture more precise information for volatilities.

Figure 5 depicts the selected longest time homogeneous interval for each time point. One can compare the selected homogeneous intervals with the LPA estimators, all of which provide evidence for the performance of the LPA. In the initial setting, we have two jumps respectively at 250, and 750. It is obvious in this figure that two jump points locate respectively round 300 and 800. Both the 5%–95% pointwise confidence

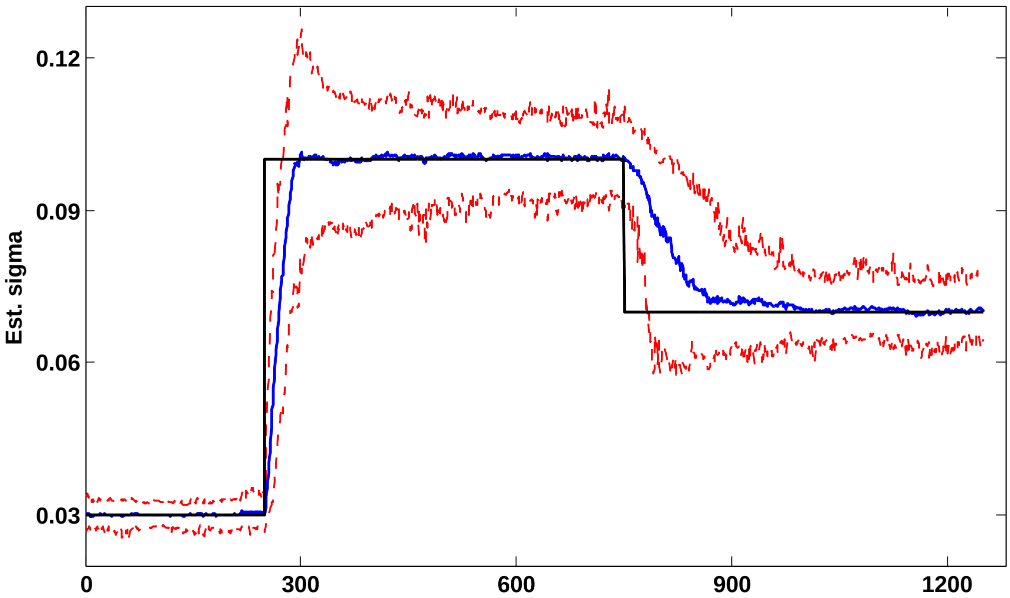

Figure 4: LPA estimator $\hat{\sigma}$ with simulated CIR paths. The dotted red lines are the 5%–95% confidence interval of $\hat{\sigma}$, the blue line is the mean of $\hat{\sigma}$, and the black line stands for true process as set in Table 1.

	$Mean$	SD	$Skewness$	$Kurtosis$
r_t	0.0319	0.0176	-0.1159	-1.4104
dr_t	-1.764×10^{-5}	0.0006	-0.7467	34.4856

Table 2: Statistical summary of three month treasury bill rate (daily data) with the period from 2 January,1998 to 13 May, 2009

intervals and the mean of the length of intervals coincide with the parameter settings.

5 Empirical Study

5.1 Data Description

We use the three month treasury bill rate from Federal Reserve Bank of St. Louis as a proxy for the short rate. It has been used frequently in the term structure literatures. The data consists of 2840 daily observations, ranging from 2 January,1998 to 13 May, 2009. The summary statistics are shown in Table 2.

The short rate and its daily change are displayed in Figure 6. Obviously, the volatil-

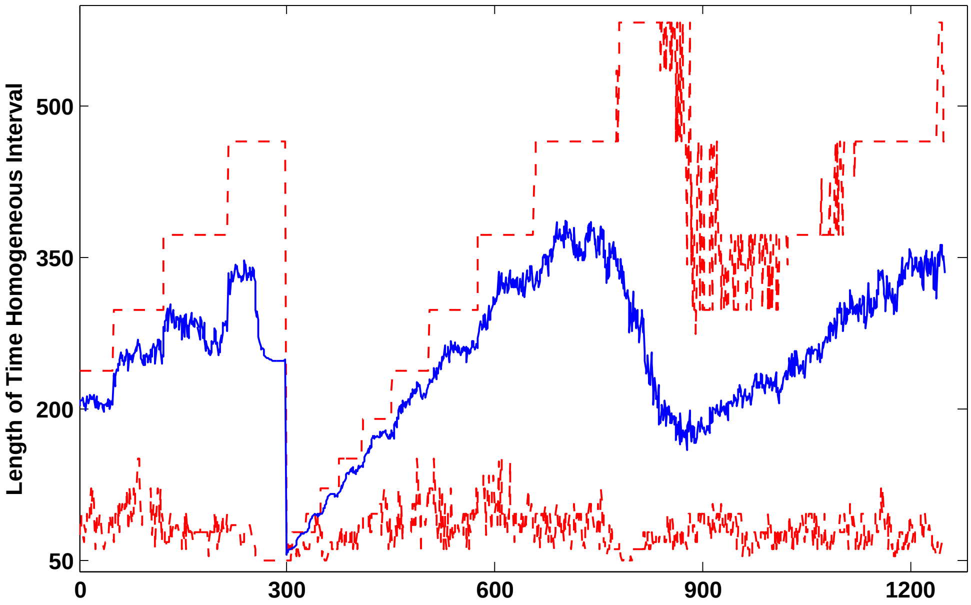

Figure 5: The length of time homogenous intervals for simulated CIR paths. The dotted red lines are the 5%–95% confidence interval, the blue lines is the mean of the estimators length of time homogeneous intervals.

ity of interest rate is changing over time. Without doubt, there are jumps and break points during the whole period; the interest rate from 1999 to 2001 is little volatile, while from mid 2007 to 2009, the volatility of interest rate is higher than that in other periods. On the basis of the phenomenon we observe from the figure that the variation of interest rate is time varying, we fit the CIR model separately with three different scenarios, the first estimation is using the whole sample, another is with the observations from the beginning of 1998 to the end of July 2007, and the last estimated period is from August 2007 to May 2009. The results are presented in Table 3. All of the three parameters differ significantly during the three different periods. For instance, the speed of mean reversion $\hat{a}$ is around 0.26 when using the whole sample, and it changes to 0.14 with the data from 1998 to 2007, and in the last period, it jumps to 3.69. Similar performance can be detected for the long run mean $\hat{b}$. Interestingly, for the volatility, it is relative low from 1998 to 2007, while it increases to 0.228 in the last period, which can be verified by Figure 6.

Firstly, we use the moving window estimation to investigate the stability of the coefficients in the CIR model. We specify three different window sizes as $l = 250$, $l = 500$, $l = 750$. Figure 7, 8 and 9 separately presents the moving window estimates $\hat{a}$, $\hat{b}$ and $\hat{\sigma}$. Quite similar performances are illustrated both in $\hat{a}$ and $\hat{b}$. One

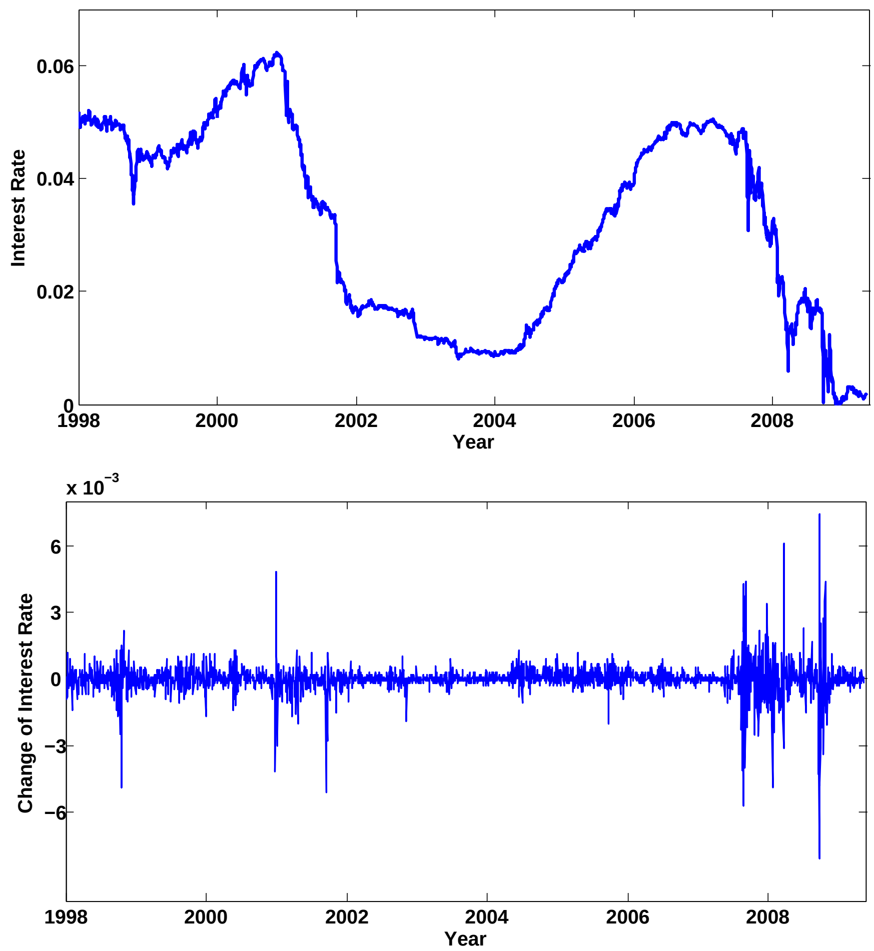

Figure 6: Three month treasure bill rate: 19980102—20090513. Top panel: Daily yields; Bottom panel: Changes of daily yields.

Sample Size	$\hat{a}$	$\hat{b}$	$\hat{\sigma}$
19980102–20090513	0.2657	0.0153	0.0944
19980102–20070731	0.1424	0.0252	0.0428
20070731–20090513	3.6792	0.0081	0.2280

Table 3: Estimated parameters of CIR model by MLE with three different time periods.

can find that Large variations exist in the process. The moving window estimator $\hat{a}$ with a very large variation is shown in Figure 7. It is not surprising that $\hat{a}$ as in the simulation is very sensitive to the data and the length of interval, even for the window size 750, it is somewhat unstable. Similarly, big jumps exist in $\hat{b}$. It can be negative at some point, and always fluctuates a lot in the different periods. However, the volatility $\hat{\sigma}$ performs in a much more stable way. It keeps almost the same value except in the last periods, where it jumps to a high volatility level.

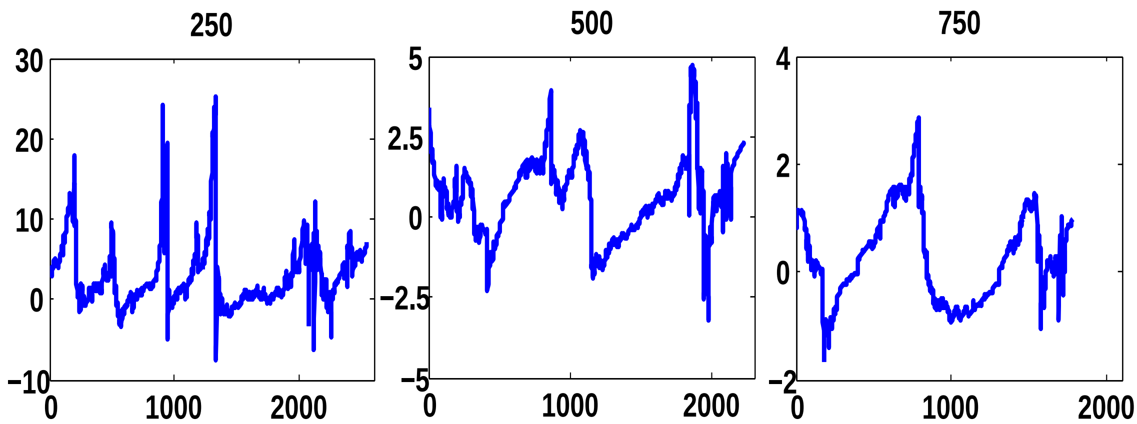

Figure 7: Moving window estimator $\hat{a}$ with window sizes 250, 500 and 750 (from left to right).

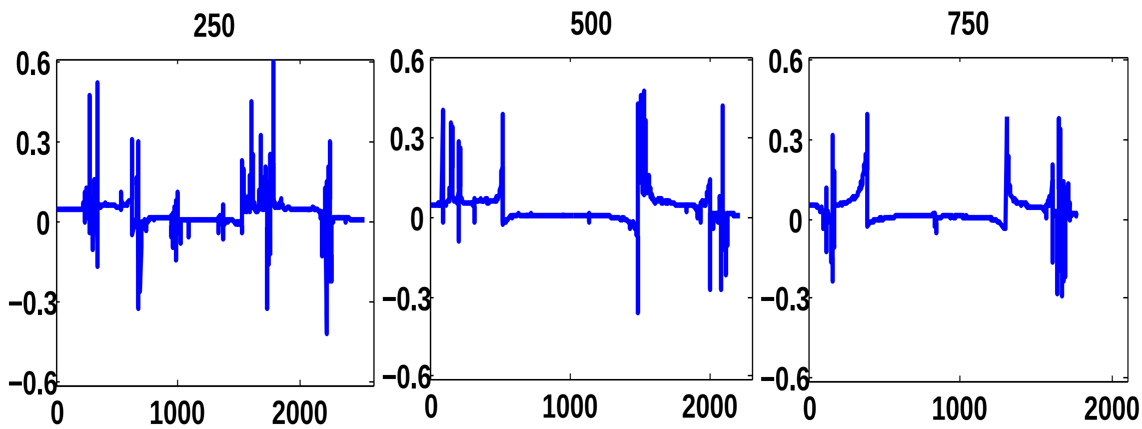

Figure 8: Moving window estimator $\hat{b}$ with window sizes 250, 500 and 750 (from left to right).

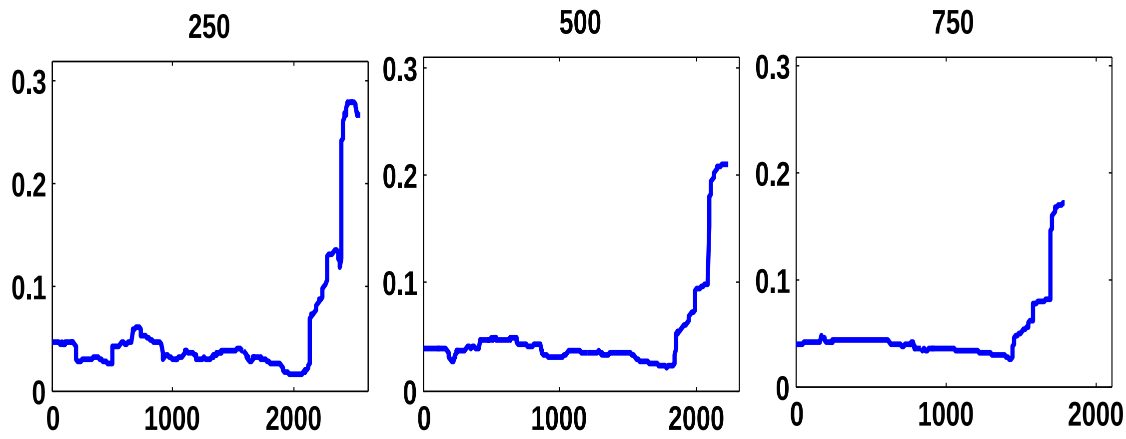

Figure 9: Moving window estimator $\hat{\sigma}$ with window sizes 250, 500 and 750 (from left to right).

For the LPA algorithm, we calculate the critical values from 500 Monte Carlo runs. We simulate the CIR path with different combinations of $\hat{a}$, $\hat{b}$, $\hat{\sigma}$ which are chosen from the estimators using different subsamples of the real data. The performance of the critical values is described in Figure 10. One can notice, the critical value is a decreasing function with respect to the log length of intervals, which is consistent with the theory mentioned above. Moreover, although we change the parameter settings for the simulation, under the null, there are not very significant differences between the critical values. We therefore choose the critical values based on the combination of the values globally estimated from data, i.e. $\theta_0^\top = (0.2657, 0.0153, 0.0944)^\top$.

The test results are shown from Figures 11 to 14. The $\hat{a}$ performs very similarly to the moving window estimators. The interest rate volatility is characterized by a fast mean reverting behaviour reflecting the impact of transient economic shocks, such as central bank announcements of base rate changes. $\hat{b}$ performs volatile in different periods, which is consistent with the behaviour of the length of selected time homogeneous interval, Figure 14. It is stable from 1999 to 2000, while the variation becomes larger in 2001 to 2003. From 2003 to 2007, it turns to be stable again, however in the last part, it reverts to a large variation again.

$\hat{\sigma}$ performs relatively stable compared with the other two CIR estimators. We again find three different regimes: from 2001 to 2003, the fluctuation of σ is increased; from mid 2007, the variance jumps to a high level, which is also reflected in the

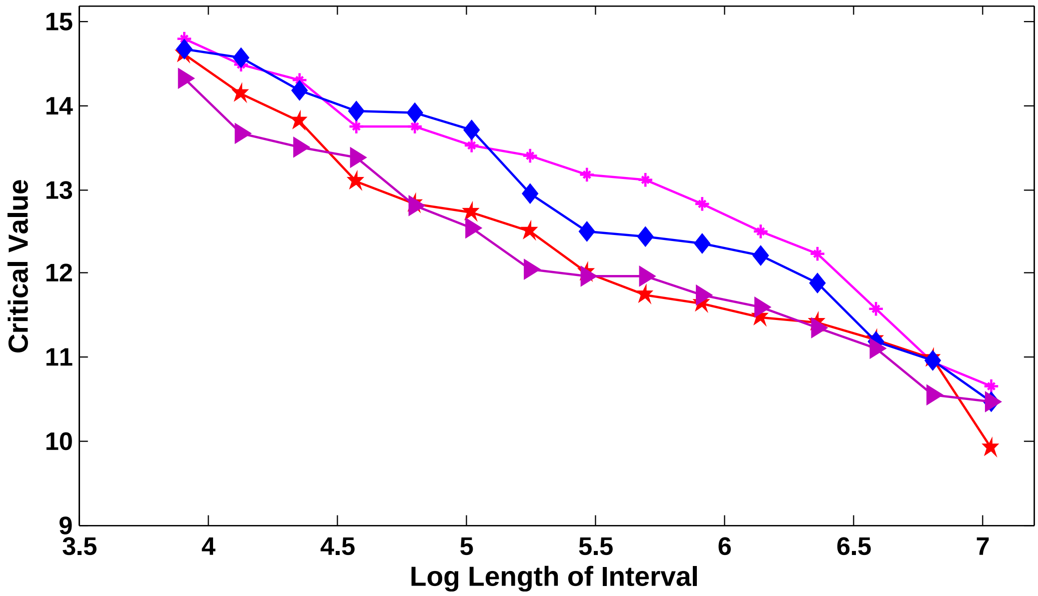

Figure 10: Critical values for four combinations of θ, changing one or two from a, b, σ with $m_0 = 40$, $K = 15$ and initial value $r_0 = 0.05$ referred by the real data.

length of the intervals $I_{\hat{k}}$, Figure 14.

Figure 14 describes the time homogeneous intervals for each time point t, here we evaluate from 1999, and we assume the first year is a time homogeneous interval. We then compare the performance of the LPA with the moving window method.

It is worth noting that the length of the selected time interval has a close relationship with the regimes of the macroeconomy. On one hand, the recession regime induces shorter homogeneous intervals, and on the other hand, the length is extended in blooming periods, where the macroeconomy is in a stable state. Let us first analyze the interest rate before 2001. In that period, the economic activity continued to expand briskly, and the variation was relatively small. We go on to compare the short rate in 2001-2003 with the selected time homogeneous interval. In this period, the US economy went into recession. It was influenced by the terrorist attack on 11 September, 2001, the stock market crash in 2002 and the war in Iraq in 2003, which induced an instable macroeconomy: increased oil prices, overstretched investment, too high productivity. All of these factors led to short selected intervals. From 2004 to 2006, the economy headed towards a stable state again. The selected intervals last longer than before. From 2007, the situation reversed, another global

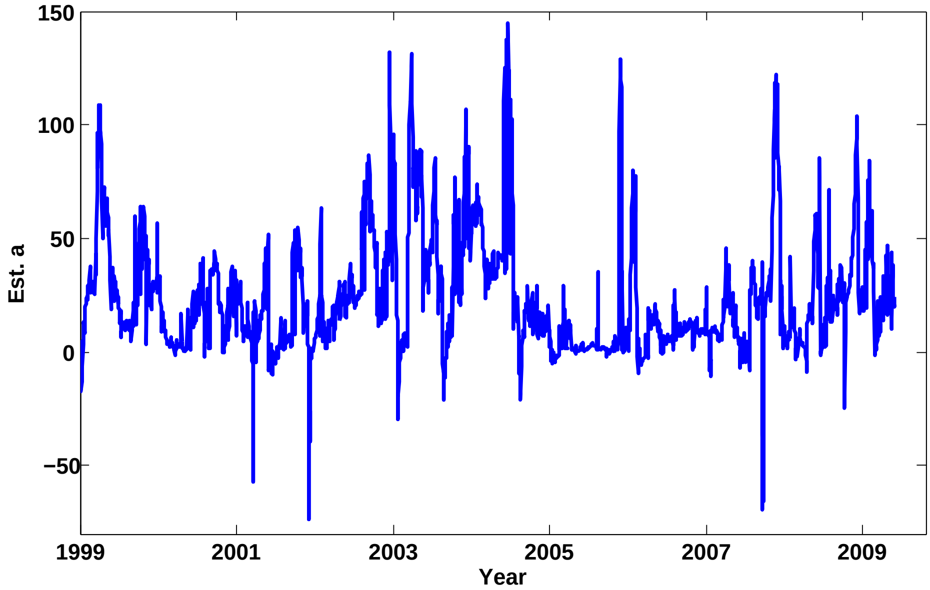

Figure 11: Estimated $\hat{a}$ for CIR model using three month treasure bill rate by LPA.

recession came. Again it can be confirmed by the shorter length of the selected intervals.

Figure 15 depicts the in sample fitting. The data is described by the black line, and the two red dashed lines stand for 10%–90% pointwise confidence intervals from the simulated data, which is the same as calculating the critical values. The blue line is the in sample fitting path with the values estimated by LPA, and the purple one is a randomly selected CIR path from the simulation. One may notice that the LPA estimated sample path matches the real data path very well, i.e. the LPA has an acceptable performance for in sample fitting. The structural break points from the fitted LPA path occur very closely to the real data path.

We further evaluate the forecasting performance of the LPA. We compare the forecasting result with the moving window estimators by means of absolute prediction errors (APE). It is defined over a prediction period horizon $\mathcal{H}$, $\mathrm{APE}(t) = \sum_{h\in\mathcal{H}} |r_{t+h} - \hat{r}_{t+h|t}|/|\mathcal{H}|$, where $\hat{r}_{t+h|t}$ represents the interest rate prediction by a particular model. Both one-day and ten-day ahead forecasting are considered. Figures 16 to 18 present the performance. In each figure, the left panel stands for the ratio from the forecasting with horizon of one day, and the right panel presents the ten days ahead forecasting. It is clear to see that the LPA performs well especially

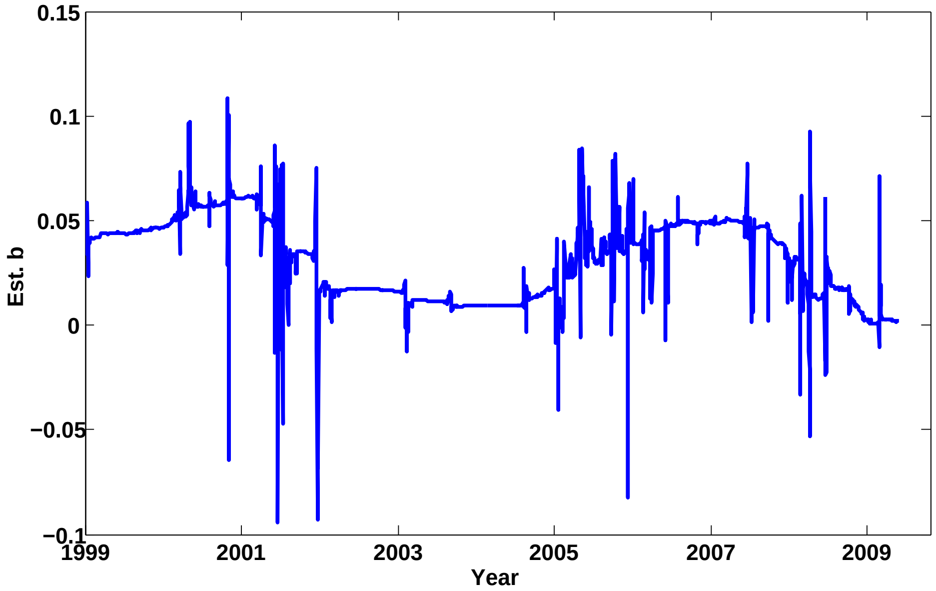

Figure 12: Estimated $\hat{b}$ for CIR model using three month treasure bill rate by LPA.

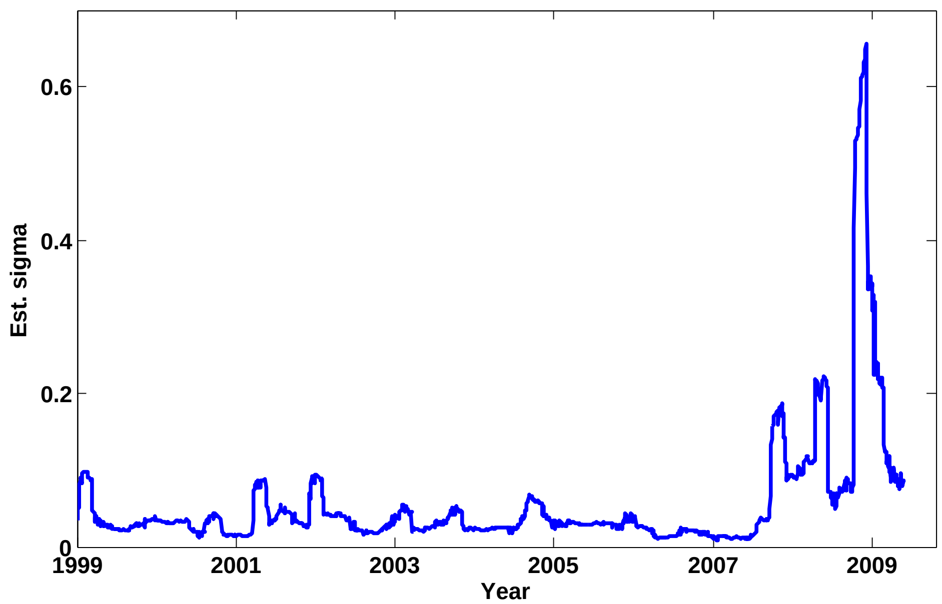

Figure 13: Estimated $\hat{\sigma}$ for CIR model using three month treasure bill rate by LPA.

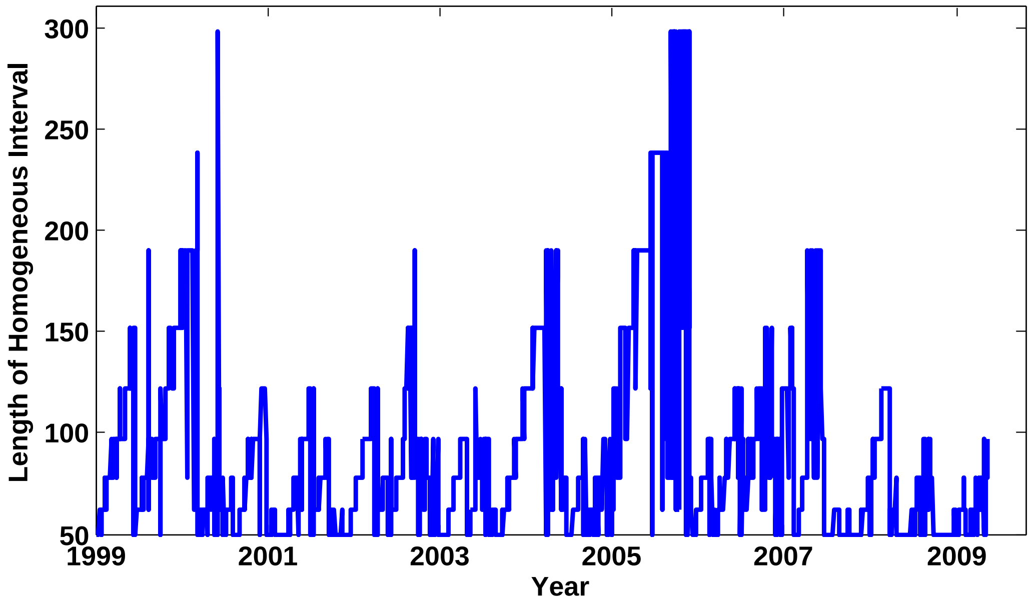

Figure 14: The selected longest time-homogeneous intervals using three month treasure bill rate with $\rho = 0.2$, and $r = 0.5$. The first reported time period is in 1999.

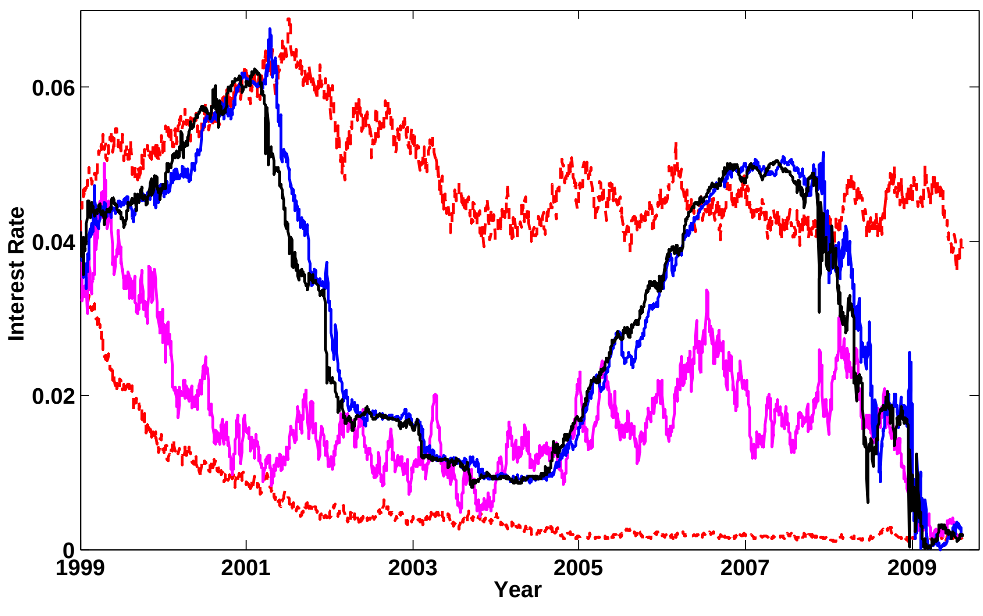

Figure 15: In-sample fitting for CIR model using three month treasure bill rate. The black line is the real data; The blue line is the fitted CIR path with the estimators by LPA; The two red lines are 10%–90% confidence intervals simulated with the global estimators; The purple line is a random selected CIR path.

in the long horizon forecasting.

First, let us consider the result from one-step ahead forecasting. One sees that, in general, the LPA is more preferable than the moving window estimation. Furthermore, as we increase the moving window size, the variation of the ratio becomes smaller, it is therefore obvious that the LPA performs relatively better, but when the economy is in an unstable state, the LPA for one step forecasting can not perform very precisely.

Next, we discuss the prediction results with the horizon of ten days (i.e. 2 weeks). The results are very interesting. In comparison with one-step forecasting, the variation becomes smaller, and the ratios are more stable. Secondly, the LPA shows a superior prediction performance. It is worthy noting that generally for ten-day ahead forecasting, the LPA outperforms the moving window estimate in the whole period. Additionally, the LPA forecasting performance improves as we compare with longer moving window estimators, because it is not reasonable to assume the parameter remains the same in a long period. The prediction is clearly better no matter if it is in the stable state or in the volatile state, which indicates the proposed LPA method shows advantages of forecasting. Additionally, we can confirm that the moving window estimations can not be valid in long horizon forecasting.

Table 4 summaries the prediction performance for the LPA and moving window (MW) estimations with the forecasting horizon of one day and ten days. We consider the mean of absolute forecasting errors (MAE) for each method. Note that for one-day ahead forecasting, there is no significant difference between the LPA and the MW, and both of their MAE are quite small. However, in ten-day ahead forecasting, the difference becomes huge. The accuracy of the MW decreases a lot compared with the LPA, especially if we increase the window size, it is more obvious.

6 Conclusion

There are both considerable statistical evidence and economic reasons to believe that the short interest rate is not stable. We apply a modern statistical method to describe the dynamics of the short rate. With the simple CIR model, and by the LPA method, we detect structural break points for the interest rate process, which is consistent with the conclusion from the existing literature that the dynamics of

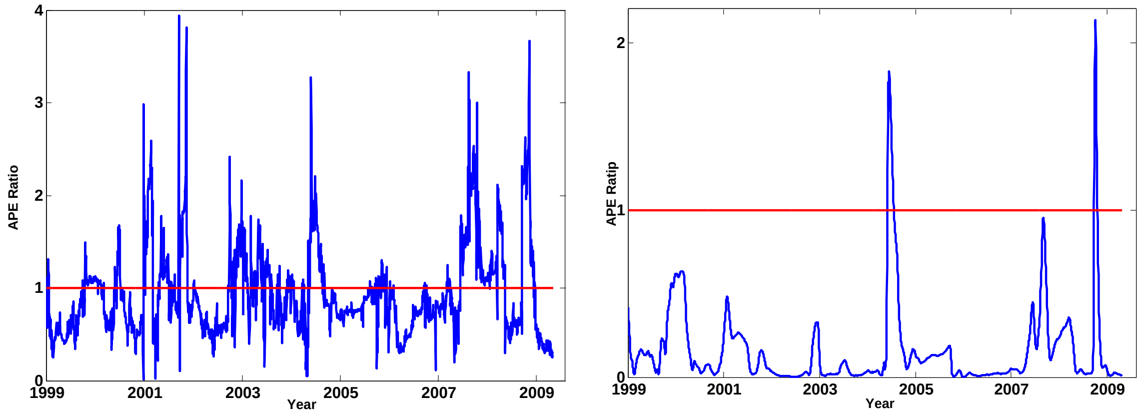

Figure 16: The ratio of the absolute prediction errors between the estimators by LPA (numerator)and moving window estimator (denominator) with window size 250. The left panel: One-day ahead forecasting; The right panel: Ten-day ahead forecasting.

Forecasting Horizon			MAE	
		$l = 250$	$l = 500$	$l = 750$
One Day	LPA	4.7409×10^{-4}	4.8516×10^{-4}	4.9649×10^{-4}
	MW	4.7851×10^{-4}	4.4181×10^{-4}	4.1681×10^{-4}
Ten Days	LPA	**0.0201**	**0.0215**	**0.0232**
	MW	0.1868	1.0032	1.8054

Table 4: The table reports the forecast evaluation criteria for one day ahead and ten days ahead forecast of the short rate based on the LPA and moving window (MW) estimation. The first column refers to the forecasting horizon. The second column represents the mean absolute forecast errors according to different moving window sizes.

interest rate is not stable. We obtain the time homogenous intervals, which is useful to explain the regime switching point. We also compare our results with the moving window estimators, and the results show that the LPA performs better in both in sample fitting and out of sample forecasting, independent of it being in a stable or unstable period.

References

Aït-Sahalia, Y. (1996). Testing continuous-time models of the spot interest rate, *Review of Financial Studies* **9:** 385-426.

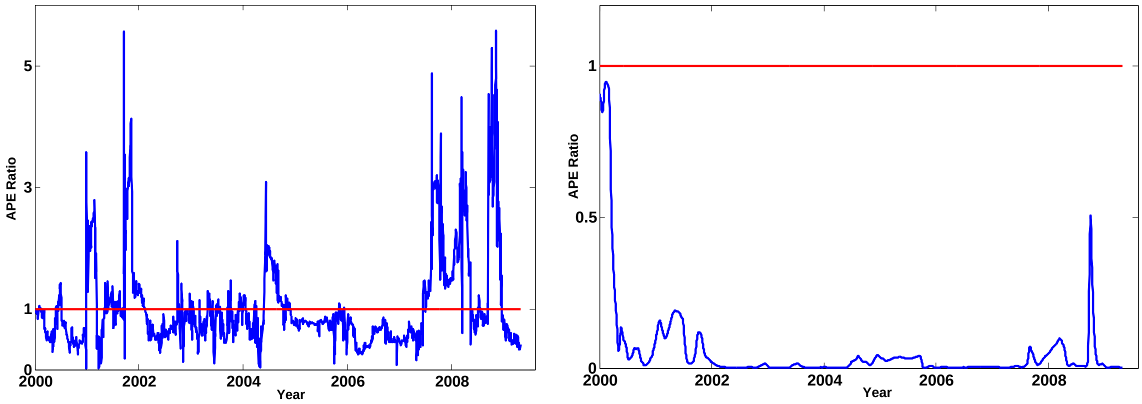

Figure 17: The ratio of the absolute prediction errors between the estimators by LPA (numerator)and moving window estimator (denominator) with window size 500. The left panel: One-day ahead forecasting; The right panel: Ten-day ahead forecasting.

Ang, A. and Bekaert, B. (2002). Regime switches in interest rates, *Journal of Business and Economic Statistics* **20:** 163-182.

Arapis, M. and Gao, J. (2006). Empirical comparisons in short-term interest rate models using nonparametric methods, *Journal of Financial Econometrics* **4:** 310-345.

Ball, C. and Torous, W. (1999). The stochastic volatility of short-term interest rates: some international evidence, *The Journal of Finance* **54(6):** 2330-2359.

Bansal, V. and Zhou, H. (2002). Term structure of interest rates with regime shifts, *The Journal of Finance* **57(5):** 1997-2043.

Brigo, D. and Mercurio, F. (2001). *Interest Rate Models - Theory and Practice*, Springer.

Black F. and Karasinski, P. (1991). Bond and option pricing when short rates are lognormal, *Financial Analysts Journal* **47:** 52-59.

Chen, Y., Härdle, W. and Pigorsch, U. (2010). Localized realized volatility modelling, *Journal of the American Statistical Association*, forthcoming.

Čížek P., Härdle, W. and Spokoiny, V. (2009). Adaptive pointwise estimation in time-inhomogeneous condtional heteroscedasticity models, *Econometric Journal* **12:** 1-25.

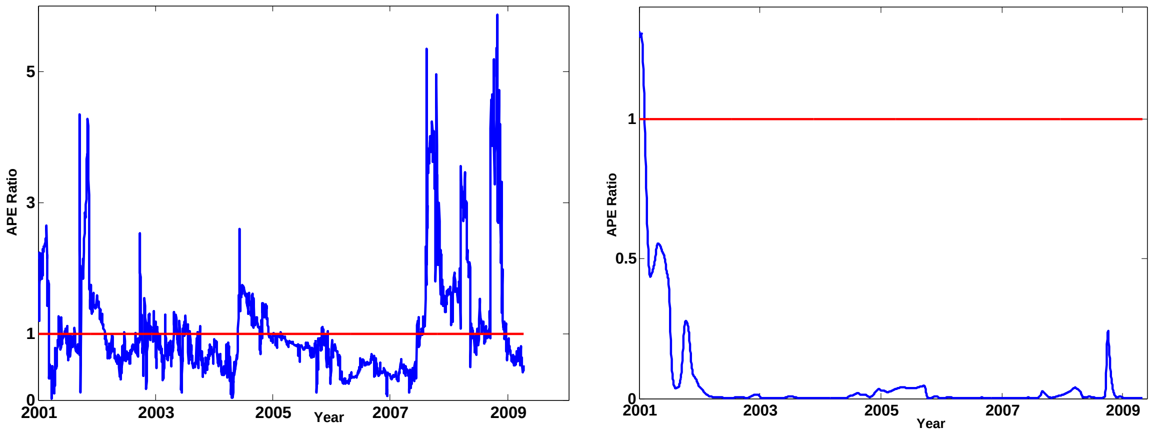

Figure 18: The ratio of the absolute prediction errors between the estimators by LPA (numerator)and moving window estimator (denominator) with window size 750. The left panel: One-day ahead forecasting; The right panel: Ten-day ahead forecasting.

Cox, C., Ingersoll, E. and Ross, A. (1985). A theory of the term structure of interest rates, *Econometrica* **53:** 385-407.

Das, S.R. (2002). The surprise element: jumps in interest rates, *Journal of Econometrics* **106:** 27-65.

Fan, J., et al (2003). Time-dependent diffusion models for term structure dynamics, *Statistica Sinica* **13:** 965-992.

Giacomini, E., Härdle, W. and Spokoiny, V. (2009). Inhomogeneous dependence modeling with time-varying copulae, *Journal of Business and Economic Statistics* **27(2):** 224-234.

Härdle, W., Franke, J. and Hafner, C. (2011). *Statistics of Financial Markets - 3rd ed*, Springer Verlag, Heidelberg, forthcoming.

Härdle, W., Okhrin, O. and Okhrin, Y. (2010). Time varying hierarchical archimedean copulae, *submitted to Journal of Applied Econometrics.*

Hull, J. and White, A. (1990). Pricing interest rate derivative securities, *Review of Financial Studies* **3(4):** 573-592.

Johannes, M.S. (2004). The statistical and economic role of jumps in continuous-time interest rate models, *Journal of Finance* **54(1):** 227-260.

Jones, M., Lamont, O. and Lumsdaine, R. (1998). Macroeconomic news and bond market volatility, *Journal of Financial Economics* **47:** 315-337.

Kladívko, K. (2007). Maximum likelihood estimation of the Cox-Ingersoll-Ross process: the matlab implementation, *working paper*.

Mercurio, D. and Spokoiny, V. (2004). Statistical inference for time-inhomogeneous volatility models, *The Annals of Statistics* **32(2):** 577-602.

Spokoiny, V. (2009). Multiscale local change point detection with applications to Value-at-Risk, *The Annals of Statistics* **37(3):** 1405-1436.

Spokoiny, V. (2010). *Local Parametric Methods in Nonparametric Estimation,* Springer Verlag, Berlin, Heidelberg, forthcoming.

Stanton, R. (1997). A nonparametric model of term structure dynamics and the market price of interest rate risk, *Journal of Finance* **52:** 1973-2002.

Vasicek, O. (1977). An equilibrium characterization of the term structure, *Journal of Financial Economics* **5:** 177-188

Zeytun, S. and Gupta, A. (2003). A comparative study of the Vasicek and the CIR model of the short rate, *Published reports of the Fraunhofer ITWM.*

SFB 649 Discussion Paper Series 2010

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Volatility Investing with Variance Swaps" by Wolfgang Karl Härdle and Elena Silyakova, January 2010.
002 "Partial Linear Quantile Regression and Bootstrap Confidence Bands" by Wolfgang Karl Härdle, Ya'acov Ritov and Song Song, January 2010.
003 "Uniform confidence bands for pricing kernels" by Wolfgang Karl Härdle, Yarema Okhrin and Weining Wang, January 2010.
004 "Bayesian Inference in a Stochastic Volatility Nelson-Siegel Model" by Nikolaus Hautsch and Fuyu Yang, January 2010.
005 "The Impact of Macroeconomic News on Quote Adjustments, Noise, and Informational Volatility" by Nikolaus Hautsch, Dieter Hess and David Veredas, January 2010.
006 "Bayesian Estimation and Model Selection in the Generalised Stochastic Unit Root Model" by Fuyu Yang and Roberto Leon-Gonzalez, January 2010.
007 "Two-sided Certification: The market for Rating Agencies" by Erik R. Fasten and Dirk Hofmann, January 2010.
008 "Characterising Equilibrium Selection in Global Games with Strategic Complementarities" by Christian Basteck, Tijmen R. Daniels and Frank Heinemann, January 2010.
009 "Predicting extreme VaR: Nonparametric quantile regression with refinements from extreme value theory" by Julia Schaumburg, February 2010.
010 "On Securitization, Market Completion and Equilibrium Risk Transfer" by Ulrich Horst, Traian A. Pirvu and Gonçalo Dos Reis, February 2010.
011 "Illiquidity and Derivative Valuation" by Ulrich Horst and Felix Naujokat, February 2010.
012 "Dynamic Systems of Social Interactions" by Ulrich Horst, February 2010.
013 "The dynamics of hourly electricity prices" by Wolfgang Karl Härdle and Stefan Trück, February 2010.
014 "Crisis? What Crisis? Currency vs. Banking in the Financial Crisis of 1931" by Albrecht Ritschl and Samad Sarferaz, February 2010.
015 "Estimation of the characteristics of a Lévy process observed at arbitrary frequency" by Johanna Kappusl and Markus Reiß, February 2010.
016 "Honey, I'll Be Working Late Tonight. The Effect of Individual Work Routines on Leisure Time Synchronization of Couples" by Juliane Scheffel, February 2010.
017 "The Impact of ICT Investments on the Relative Demand for High-Medium-, and Low-Skilled Workers: Industry versus Country Analysis" by Dorothee Schneider, February 2010.
018 "Time varying Hierarchical Archimedean Copulae" by Wolfgang Karl Härdle, Ostap Okhrin and Yarema Okhrin, February 2010.
019 "Monetary Transmission Right from the Start: The (Dis)Connection Between the Money Market and the ECB's Main Refinancing Rates" by Puriya Abbassi and Dieter Nautz, March 2010.
020 "Aggregate Hazard Function in Price-Setting: A Bayesian Analysis Using Macro Data" by Fang Yao, March 2010.
021 "Nonparametric Estimation of Risk-Neutral Densities" by Maria Grith, Wolfgang Karl Härdle and Melanie Schienle, March 2010.

SFB 649 Discussion Paper Series 2010

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

022 "Fitting high-dimensional Copulae to Data" by Ostap Okhrin, April 2010.
023 "The (In)stability of Money Demand in the Euro Area: Lessons from a Cross-Country Analysis" by Dieter Nautz and Ulrike Rondorf, April 2010.
024 "The optimal industry structure in a vertically related market" by Raffaele Fiocco, April 2010.
025 "Herding of Institutional Traders" by Stephanie Kremer, April 2010.
026 "Non-Gaussian Component Analysis: New Ideas, New Proofs, New Applications" by Vladimir Panov, May 2010.
027 "Liquidity and Capital Requirements and the Probability of Bank Failure" by Philipp Johann König, May 2010.
028 "Social Relationships and Trust" by Christine Binzel and Dietmar Fehr, May 2010.
029 "Adaptive Interest Rate Modelling" by Mengmeng Guo and Wolfgang Karl Härdle, May 2010.